

## Compressão e Percepção

- FINISHED FILES ARE THE RESULT OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF YEARS.

Quantos F's tem a frase

# Compressão e Percepção



Quantos F's tem a frase

## Compressão e Percepção

- FINISHED FILES ARE THE RESULT OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF YEARS.

## Compressão e Percepção

- FINISHED FILES ARE THE RESULT OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF YEARS.
- Par a smana não há auas

## Compressão e Percepção

- FINISHED FILES ARE THE RESULT OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF YEARS.
- Par a smana não há auas
- A falta de mao de obra quaificada nas empersas portugeusas proibe a economai de ...

# Compressão e Percepção

- Imagem: Tif (786kbytes)



- Jpeg (37kbytes ) - “quality” 75%)



compressão 1:20

- Audio: 🗣️ wav (37Mbytes)  
compressão 1:11

mp🗣️(3.3Mbytes) – 128kbps

- Video: 1'26'' (352x288) + audio  
ficheiro mpg 8,075,264 bytes  
compressão ?

# Compressão de sinais multimédia

- Compressão de sinais:
  - Áudio (música, fala)
  - Imagem
  - Vídeo
- Compressão com e sem perdas (lossy and lossless)

Baseia-se na capacidade de suprimir partes dos sinais mantendo as mesmas características perceptíveis ao homem.

A compressão baseia-se em redundâncias estatísticas, temporais, espaciais, espectrais e estruturais.

# Espectro Visível

- A luz observada é função da luz incidente no objeto e da refletividade ou transmissividade do objeto

$$I(\lambda) = \rho(\lambda)L(\lambda)$$

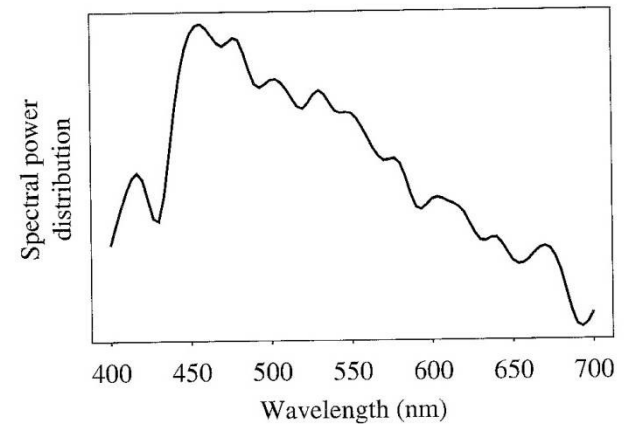
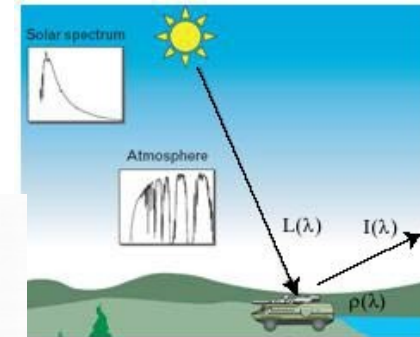
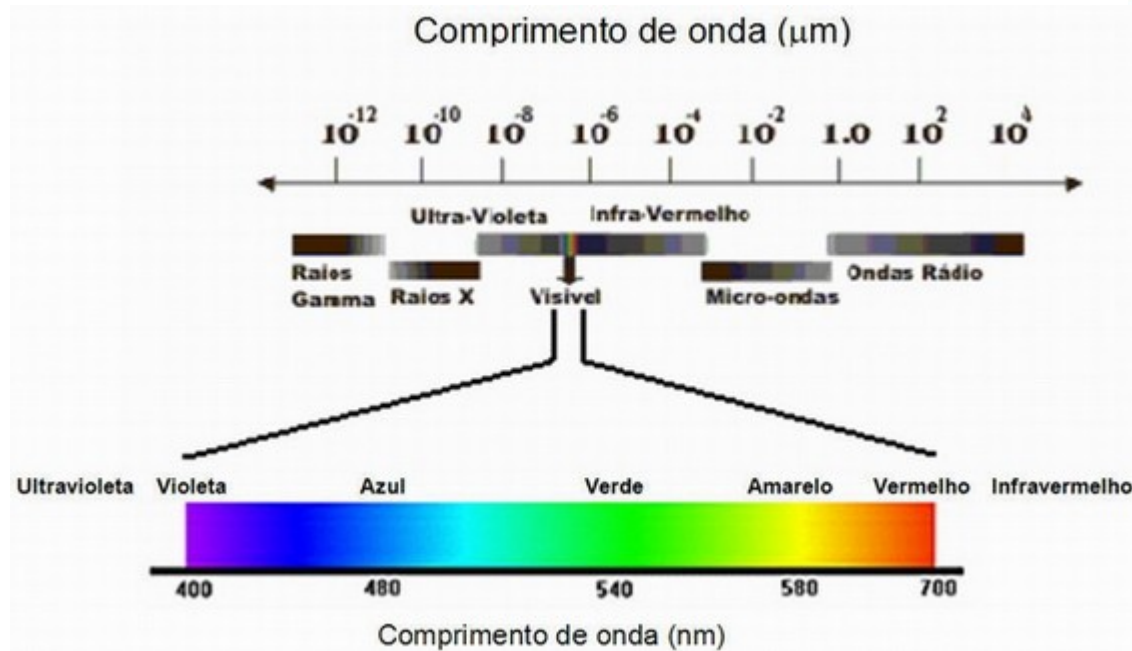


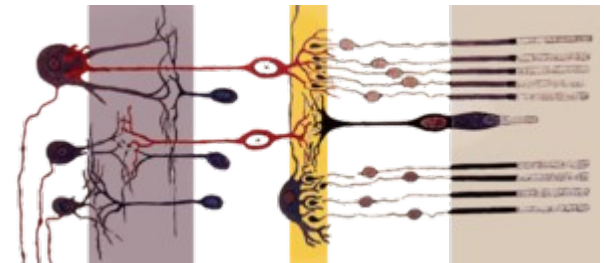
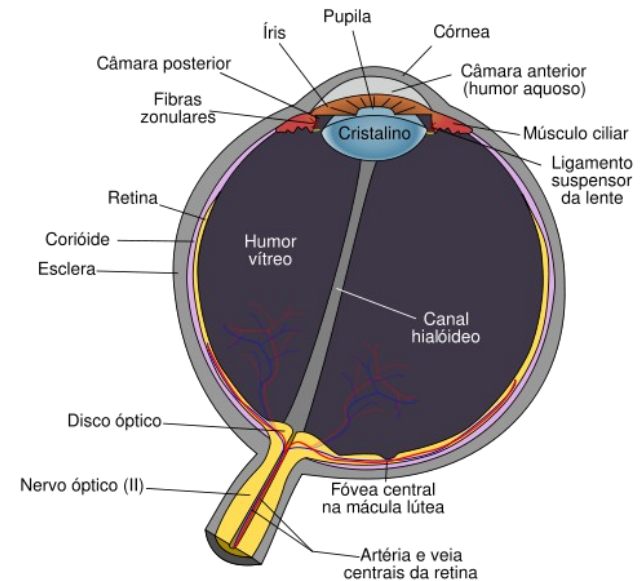
FIGURE 4.2: Spectral power distribution of daylight.

- Nota: há cores que não têm apenas um comprimento de onda (ex: cor purpura)



# Sistema Visual Humano

- Aquisição de ondas electromagnéticas é responsável pela 70% da informação adquirida pelo 5 sentidos.
- Íris: diafragma que controla a abertura da pupila
- Cristalino: Funciona como lente
- cones (região fovea) sensíveis à cor (sensação cromática)
- bastonetes sensíveis à luminosidade (sensação acromática)
- Nota: Á noite não vemos a cor!

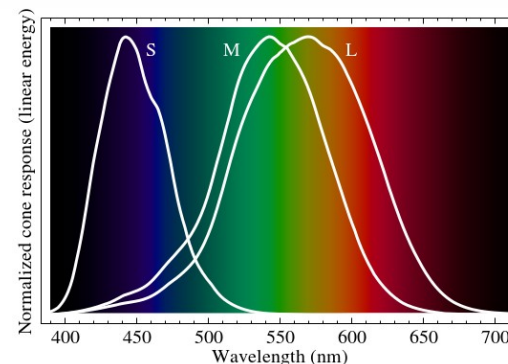
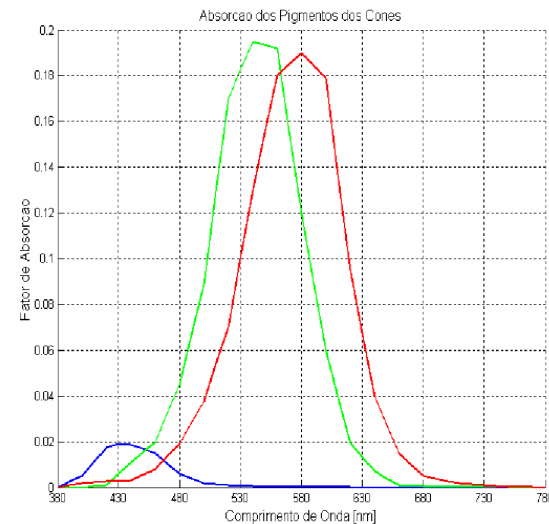
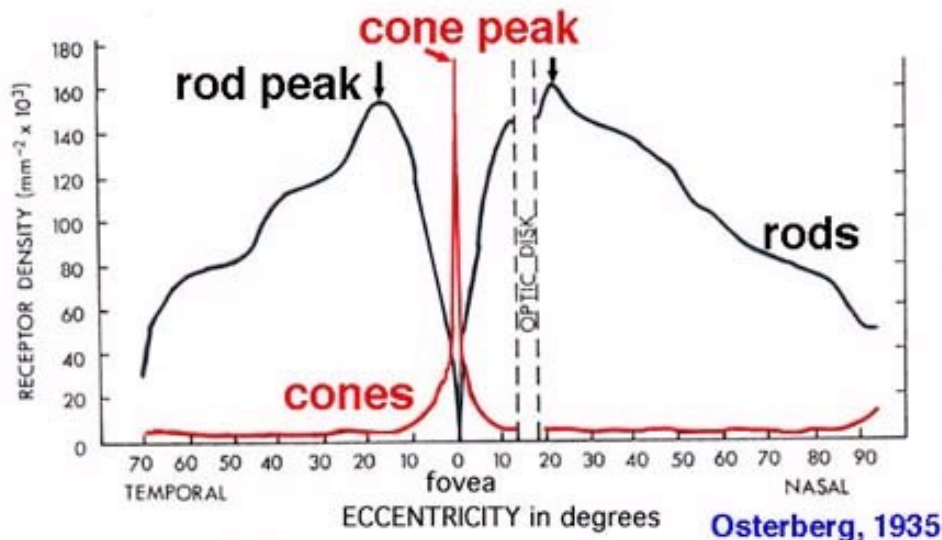


# Sistema Visual Humano (percepção cromática)

- Região Fovea: 1-2° (1mm<sup>2</sup>)
- Cones abundantes só em cerca de 5° (4.5-6 milhões)
- Bastonetes em cerca de 20° (90-120 milhões)
- Bastonetes 1000 vezes mais sensíveis que cones (basta um único fóton)

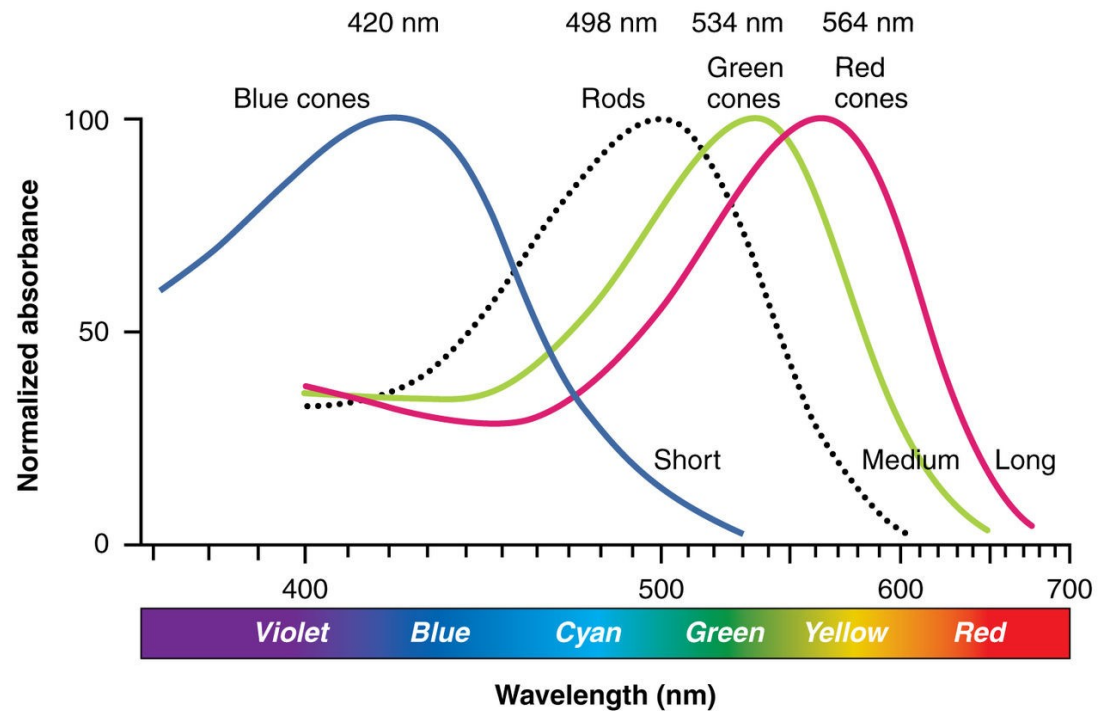
## ■ Cones

- Verde (540 nm): 33%
- Vermelho (580 nm): 65%
- Azul (440 nm) : 2%



# Sistema Visual Humano (percepção cromática)

- A visão humana consegue distinguir:
  - $10^6$ - $10^7$  cores;
  - $\gg 10^7$  diferentes brilhos

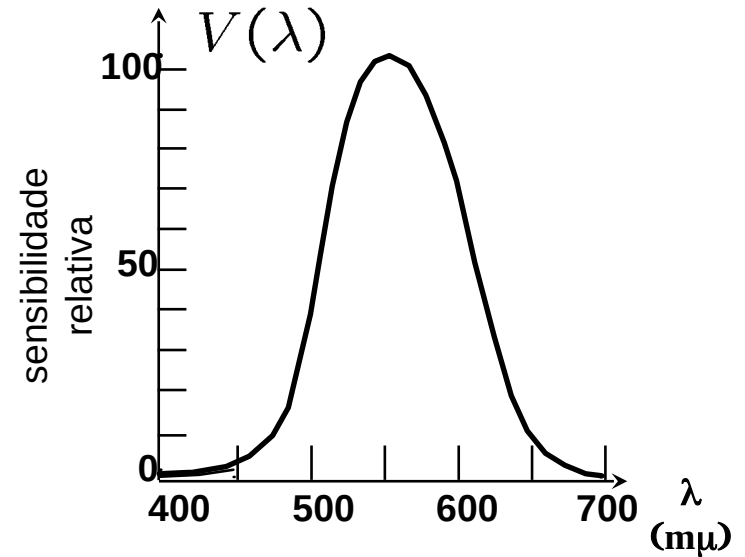


# Sistema Visual Humano (pecepção acromática)

- Luminância (intensidade luminosa)

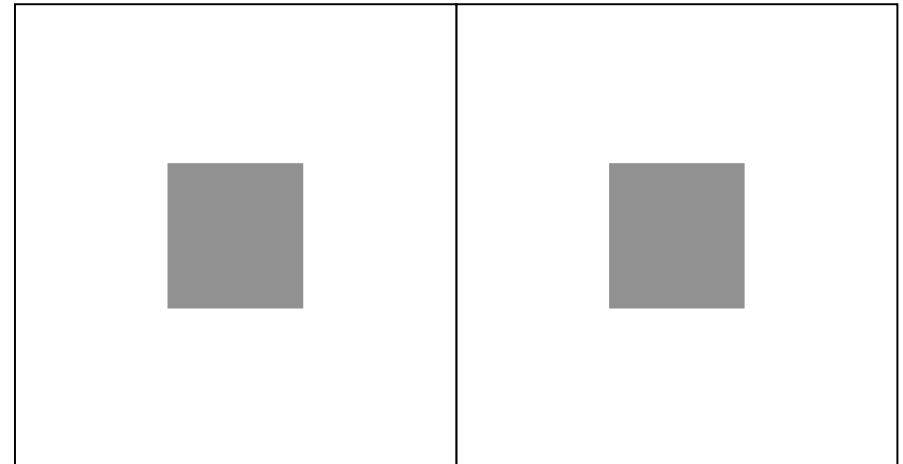
$$L = \int_0^{\infty} I(\lambda)V(\lambda)d\lambda \text{ cd/m}^2$$

- Função de eficiência luminosa (na zona fovea)



- **Nota:** A percepção da luminância de um objecto depende do contraste com o fundo.

- O olho humano consegue distinguir centenas de níveis de cinzento

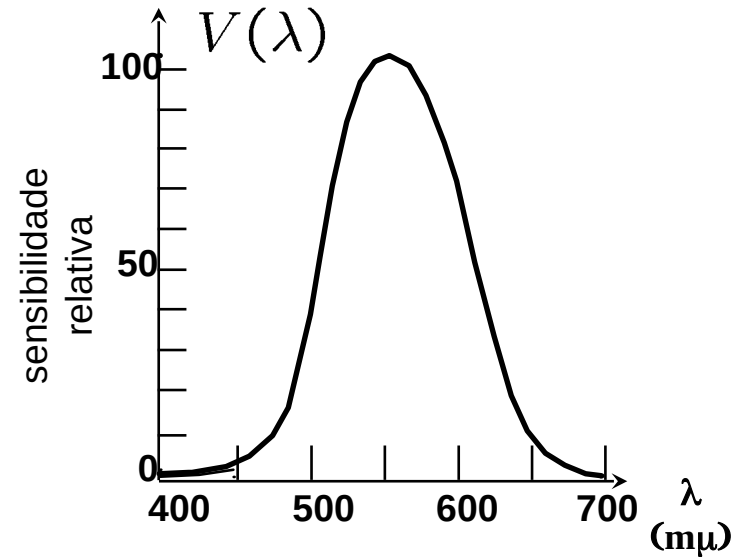


# Sistema Visual Humano (percepção acromática)

- Luminância (intensidade luminosa)

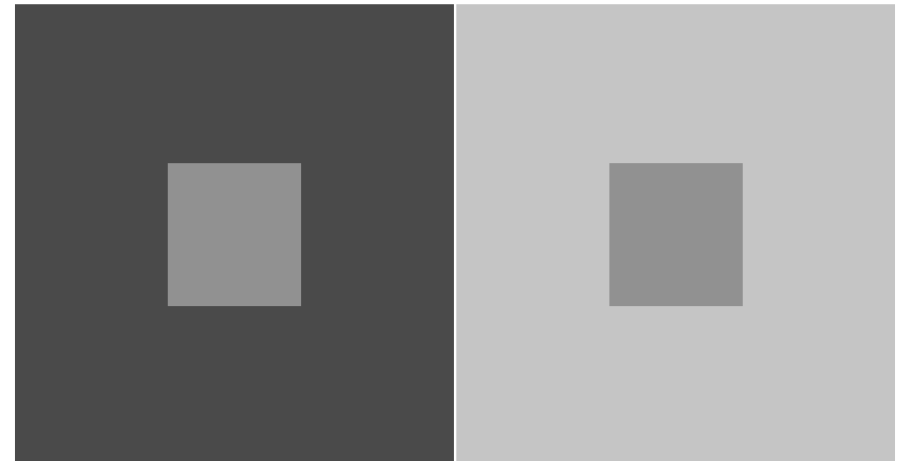
$$L = \int_0^{\infty} I(\lambda)V(\lambda)d\lambda \text{ cd/m}^2$$

- Função de eficiência luminosa (na zona fovea)



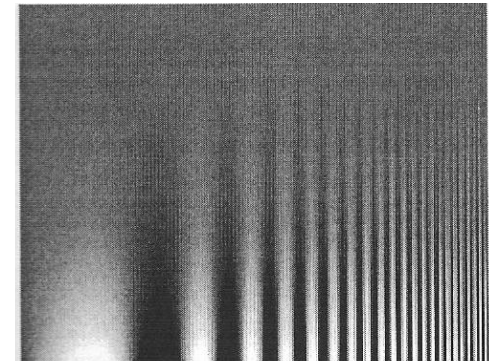
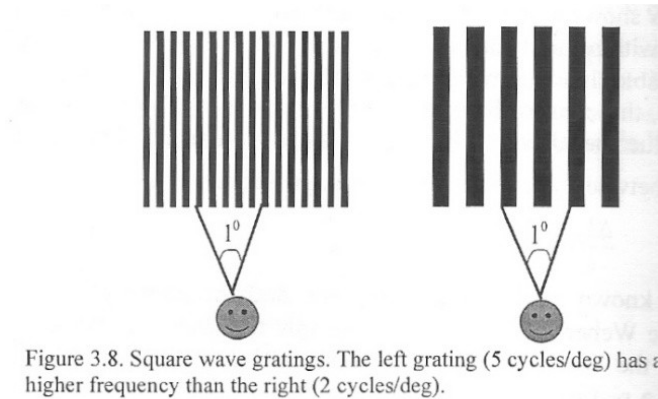
- **Nota:** A percepção da luminância de um objecto depende do contraste com o fundo.

- O olho humano consegue distinguir centenas de níveis de cinzento



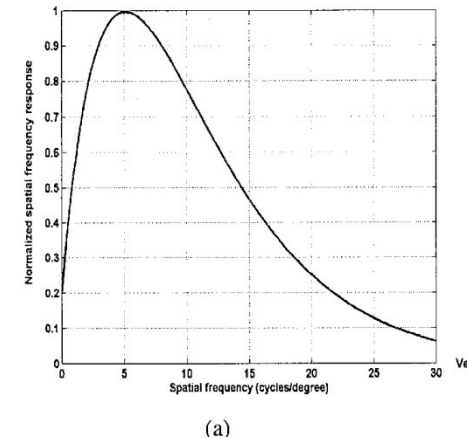
# Sistema Visual Humano (percepção acromática)

- Resposta em frequência espacial



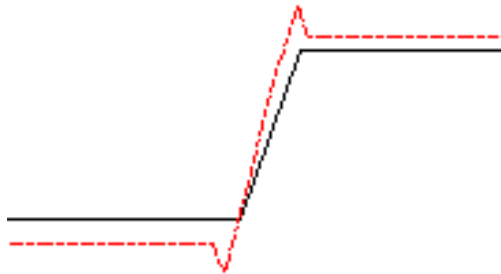
- **Nota:**

- O olho tem mais acuidade entre 2 a 10 ciclos por grau.
- A sensibilidade do olho é menor em direcções diagonais.



## Sistema Visual Humano (percepção acromática)

- Efeito da banda de Mach
  - A percepção da luminância não é monotónica
  - Sensibilidade aos contornos



## Acuidade visual

- Ângulo de visão
- Resolução espacial
  - O mais pequeno detalhe perceptível à visão humana é cerca de 1minuto =  $1/60^\circ$

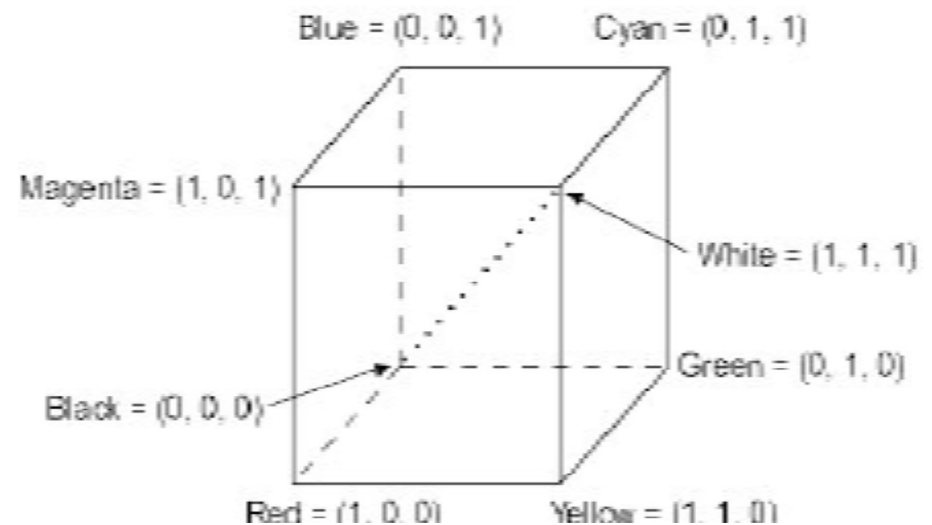
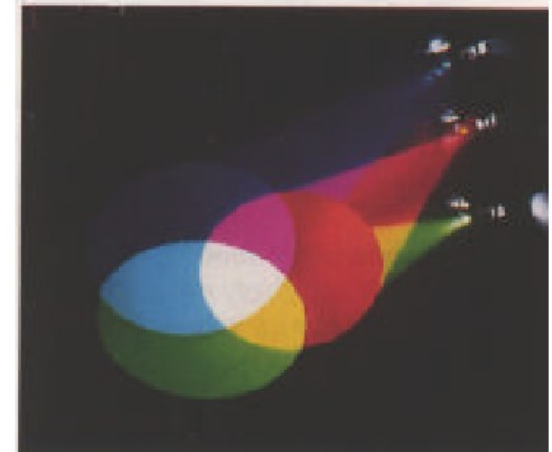
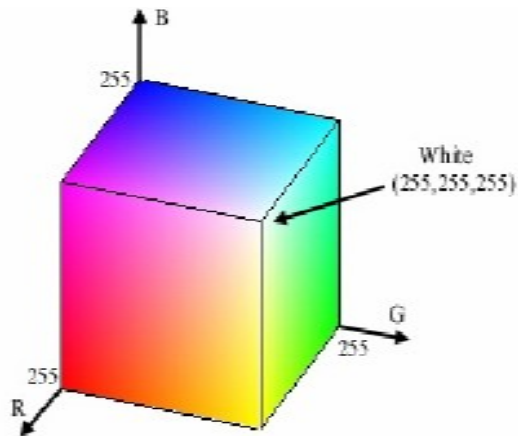
## Propriedades temporais da visão

- Persistência da visão:
  - Várias frames amostradas a uma determinada velocidade parece um movimento contínuo
  - O efeito “flicker ” é tanto maior quanto maior for a intensidade luminosa
  - Este efeito é desprezável quando a frequência for superior a 50/60hz.



# Representação de Côr em Imagem

- Sistema RGB
  - Red Green Blue
  - Sistema aditivo na mistura de côr
  - Usado nos monitores



# Representação de Côr em Imagem

- Sistema CMY
  - Cian Magenta Yellow
  - sistema subtrativo na mistura de côr
  - Conversão CMY – RGB

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



- Sistema CMYK
  - K – preto (melhora o contraste e poupa tinta)
  - Usado nas impressoras
  - Conversão CMY - CMYK

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} C - K \\ M - K \\ Y - K \end{bmatrix}$$

$$K = \min\{C, M, Y\}$$

# Representação de Côr em Imagem

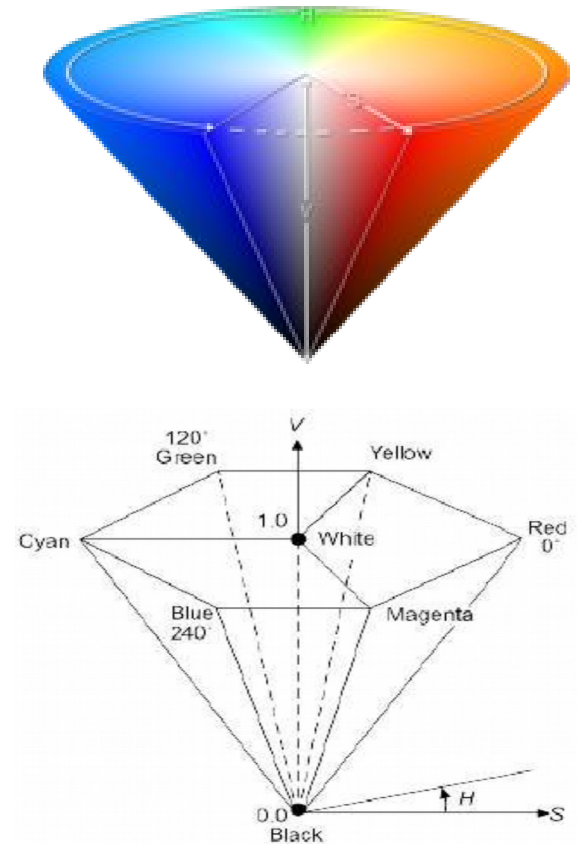
- Sistema HSV
  - Hue (cor dominante)  
varia entre 0 e 360°  
180° entre cores complementares
  - Saturation (puricidade)  
varia entre 0 e 1 (1-saturado)
  - Value (luminância)  
varia entre 0 e 1 (1-brilhante)

Branco:  $S=0$   $V=1$

Cinzentos:  $S=0$ ,  $V[0\ 1]$

Preto:  $V=0$

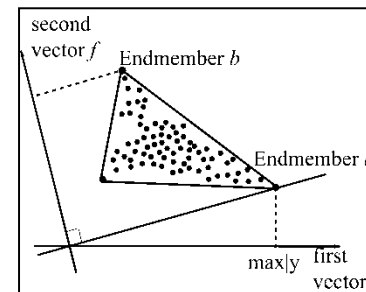
Cor pura:  $S=1$   $V=1$



# Representação de Imagens

- Tipos de ficheiros:
  - BMP, GIF, TIFF, JPEG, PS, EPS, PDF, etc
  - Há formatos que fazem compressão (com e sem perdas)

- Resolução espacial da imagem:
  - “aspect ratio” - 4:3 ou 16:9
  - (640x480) (1024x768) ou (1280x720) (1920x1080)



- Representação cor
  - Monocromática: 1bit/pixel (linhas x colunas x 1)



# Representação de Imagens

- Tipos de ficheiros:
  - BMP, GIF, TIFF, JPEG, PS, EPS, PDF, etc
  - Há formatos que fazem compressão (com e sem perdas)
- Resolução espacial da imagem:
  - “aspect ratio” - 4:3 ou 16:9
  - (640x480) (1024x768) ou (1280x720) (1920x1080)
- Representação cor
  - Monocromática: 1bit/pixel
  - Níveis de cinzento: 8bit/pixel (linhas x colunas x 8)



# Representação de Imagens

- Tipos de ficheiros:
  - BMP, GIF, TIFF, JPEG, PS, EPS, PDF, etc
  - Há formatos que fazem compressão (com e sem perdas)
- Resolução espacial da imagem:
  - “aspect ratio” - 4:3 ou 16:9
  - (640x480) (1024x768) ou (1280x720) (1920x1080)
- Representação côr
  - Monocromática: 1bit/pixel
  - Níveis de cinzento: 8bit/pixel
  - Full color: 24 bit/pixel (8 bit cada côr)  
(linhas x colunas x 8 x 3)  
Suporta 256x256x256 cores!



# Representação de Imagens

- Tipos de ficheiros:
  - BMP, GIF, TIFF, JPEG, PS, EPS, PDF, etc
  - Há formatos que fazem compressão (com e sem perdas)
- Resolução espacial da imagem:
  - “aspect ratio” - 4:3 ou 16:9
  - (640x480) (1024x768) ou (1280x720) (1920x1080)
- Representação côr
  - Monocromática: 1bit/pixel
  - Níveis de cinzento: 8bit/pixel
  - Full color: 24 bit/pixel
  - Mapeamento de côr: 8 bit/pixel ou seja 256 cores  
Usa uma “palette” de cor

Como construir a palette?



# Dithering

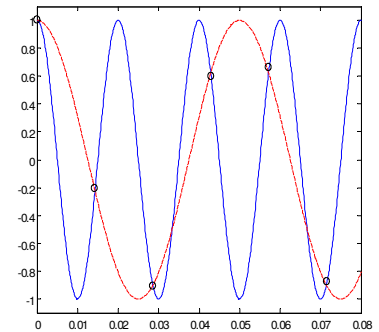
- Compromisso entre resolução de cor e resolução espacial
- A ilusão criada deve-se ao facto que a visão humana integra a intensidade em pequenas regiões da imagem
- Melhor que quantificar com 2 níveis
- Diferentes técnicas:
  - Ordered
  - Floyd-Steinberg (error diffusion)
  - Bayer
  - Halftone
  - Random





# Digitalização de sinais - PCM

- Amostragem e Quantificação
- Amostragem
  - Frequência de amostragem ( $f_s > 2f_{max}$ )
  - Aliasing
    - Em audio (1kHz amostado a)  
8kHz                      1.5kHz



- Em Imagens



# Medidas de qualidade da imagem

- Subjetivas

ou

- objetivas

$$\text{SNR(dBs)} = 10 \log_{10} \left( \frac{\sum_l \sum_c I_{\text{org}}[l, c]^2}{\sum_l \sum_c (I_{\text{org}}[l, c] - I_{\text{app}}[l, c])^2} \right)$$

$$\text{PSNR(dBs)} = 10 \log_{10} \left( \frac{\sum_l \sum_c 255^2}{\sum_l \sum_c (I_{\text{org}}[l, c] - I_{\text{app}}[l, c])^2} \right)$$

$$\text{MSE} = \sum_l \sum_c (I_{\text{org}}[l, c] - I_{\text{app}}[l, c])^2$$

$$\text{MAE} = \sum_l \sum_c |I_{\text{org}}[l, c] - I_{\text{app}}[l, c]|$$

## Modelos de Côr em Vídeo

- Transformação YCbCr – RGB

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168 & -0.331 & 0.5 \\ 0.5 & -0.418 & -0.08 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

- Na recomendação ITU-R BT.601-4:

- Codificação deve ser a 8 bit;

- Y varia entre 235 e 16

$$Y = 219(0.299R + 0.587G + 0.114B) + 16$$

- Cb,Cr variam entre 16 e 240

$$C_b = 112(B - Y)/0.866 + 128$$

$$C_r = 112(R - Y)/0.701 + 128$$

# Vídeo Digital

- Vantagens
  - Qualidade de imagem (resol. espacial, frame rate, aspect ratio)
  - Qualidade do som (4-6 canais)
  - Processado num PC
  - Encriptação
  - Melhor SNR
  - Fácil armazenamento
  - Fácil acesso para edição
- Desvantagem
  - Requer uma largura de banda elevada
  - Requer muito espaço de armazenamento
- Compromisso
  - **Compressão**
- Exemplo imagem 1920\*1080
  - Frame rate 30Hz
  - 24bit por pixel
  - 1492Mbit/seg = **186Mbytes/seg**

# Vídeo Digital

- Dado que o video é uma sequência de imagens
  - Pode-se comprimir cada imagem (intraframe)  
e
  - Pode-se tirar partido da redundância temporal  
e do sistema visual que é menos sensível a objectos em movimento  
(compressão interframe)

- Dado que o sistema da visão humana é menos sensível à cor que a intensidade a quantificação de cor pode ser mais pobre.

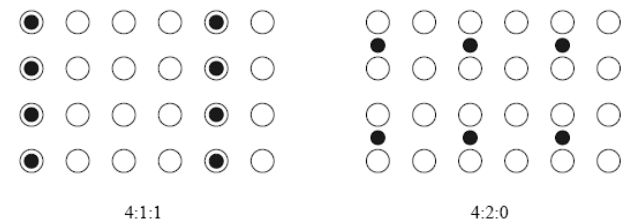
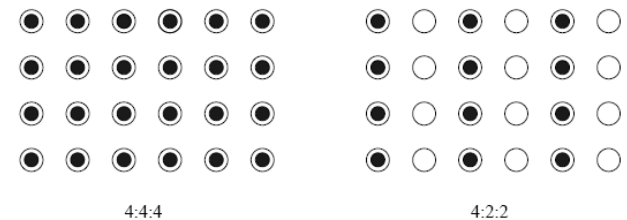
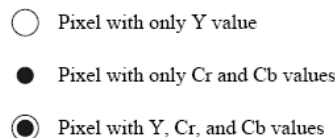
- Chroma subsampling

4:4:4 - sem subamostragem

4:2:2 - amost. hor. Cb Cr factor 2

4:1:1 - amost. hor. Cb Cr factor 4

4:2:0 - amost. hor. E vert. factor 2

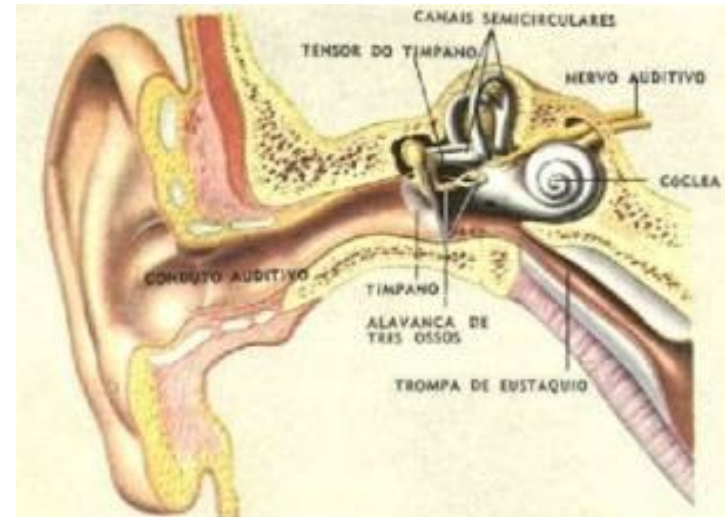


## Conceitos básicos do som

- As ondas sonoras são ondas longitudinais (mecânicas ) produzidas por deformações provocadas pela diferença de pressão num meio elástico qualquer (gasoso,sólido, liquido. ex: ar) precisando deste meio para se propagar;
- Propagam-se a uma velocidade de 344m/s (no ar a uma temp. 20°);
- Estas ondas gozam das mesmas propriedades que outras ondas mecânicas (refracção, difracção, reflecção, interferência, etc);
- Tom: onda sinusoidal, tem uma só frequência;
- Sons mais ricos contêm mais que um tom;
- Instrumentos musicais têm uma frequência fundamental e harmónicas desta.

# Sistema Auditivo Humano

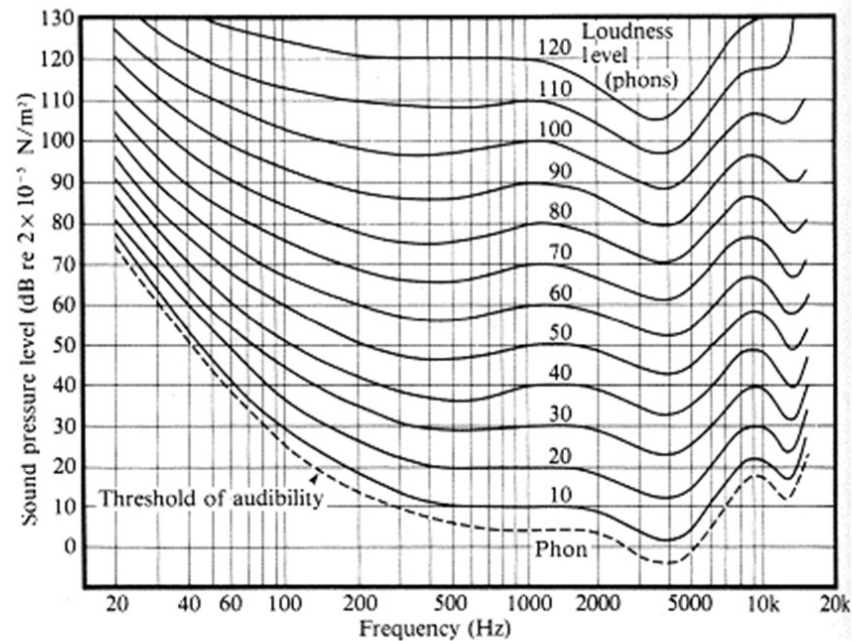
- Ouvido externo:
  - Pavilhão auditivo (orelha)
    - Colecta e filtra o som
  - Conduto auditivo Externo
    - Transmite som para o tímpano
    - Câmara de ressonância
- Ouvido médio
  - Ossículos (martelo, bigorna, estribo)
    - Conduz a vibração do tímpano ao ouvido interno
- Ouvido interno
  - Cóclea ( tem líquido no seu interior) este activa o órgão de Corti (transmite impulsos nervosos)
  - Trompa Eustáquio (equilibra a pressão interior/exterior)
  - Aparato vestibular – Equilíbrio e sistema de visão



# Sistema Auditivo Humano (percepção)

- Ouvimos o som de qualquer direcção, mas temos dificuldade em saber essa direcção
- Altamente sensível a defeitos e a sinais atractivos
- Ignora defeitos graves em sinais assumidos como irrelevantes
- Sensível a interrupções no som.
- Adaptativo
- Não-linear
  - Resposta na frequência
    - Banda 20Hz-20Khz
  - maior sensibilidade na banda 500-4000Hz

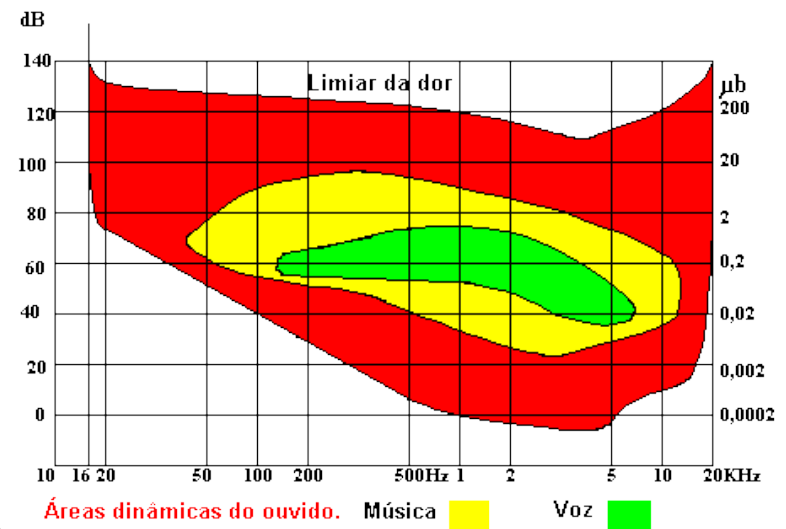
## ■ curvas de Fletcher-Munson





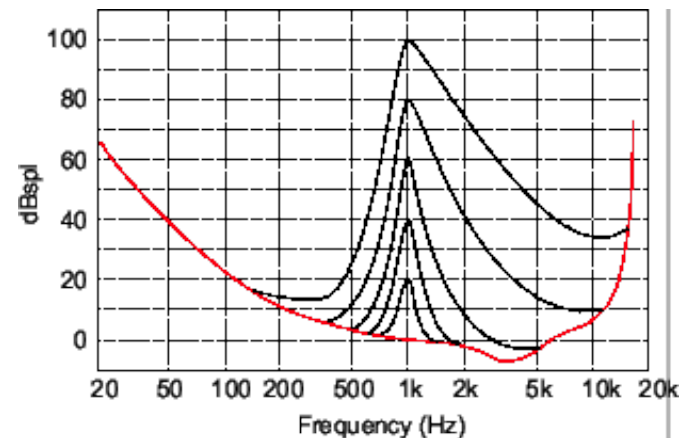
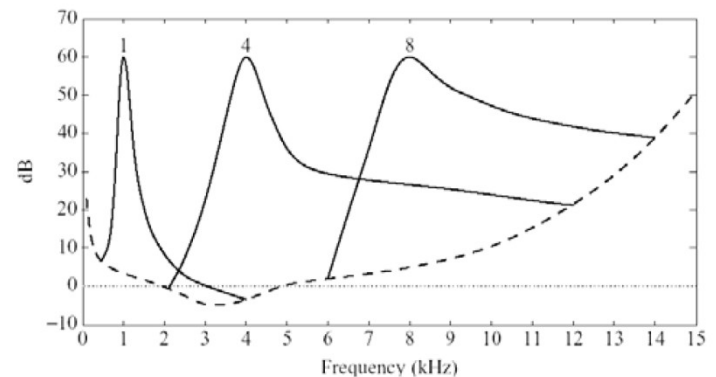
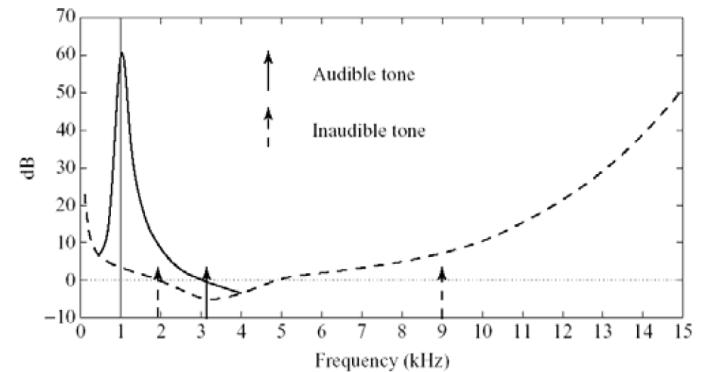
# Sistema Auditivo Humano (percepção)

- 160 dB – Lesão no tímpano
- 140 dB – Limiar da dor
- 130 dB – Barulho do avião
- 120 dB – Limiar de desconforto
- 100 dB – Concerto Rock
- 70 dB – Rua com tráfego
- 60 dB – conversa comum
- 35 dB – casa comum
- 0 db – silêncio (limiar de audição)



# Efeito de máscara na frequência

- Um tom grave pode mascarar outro mais agudo
- O contrário não é verdade
- Quanto maior a frequência  
Maior a gama de frequências afectada
- Se aumentar a potência da máscara, maior a banda afectada.



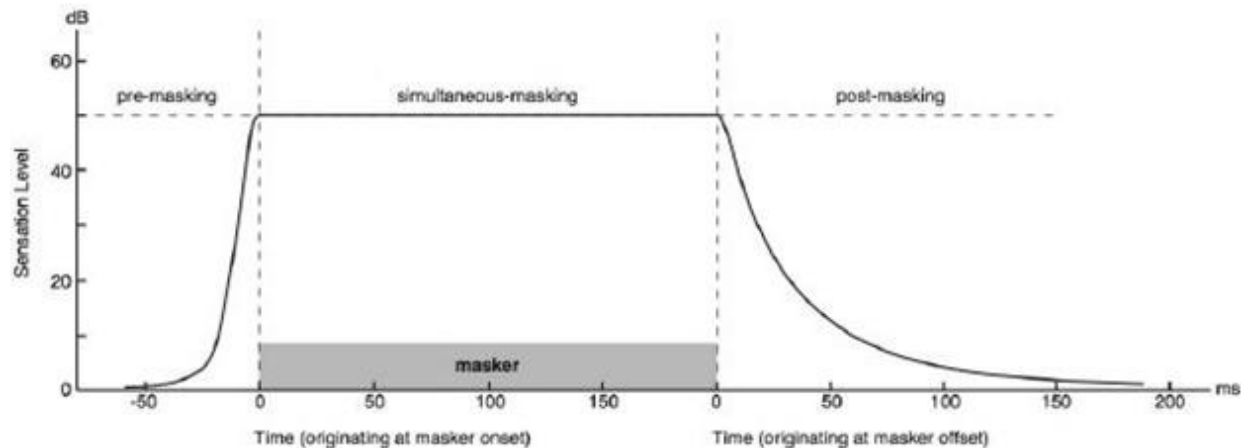
## Bandas críticas

- O sistema auditivo processa o som em sub-bandas chamadas Bandas Críticas (banco de filtros sobrepostos situados na coclea)
- São filtros não-lineares (dependem da intensidade) e as suas larguras de banda aumentam com a frequência
- Em cada sub-banda o som é analisado independentemente de outra sub-banda
- No estudo de Fletcher (1940) estas bandas representam a máscara no qual o sistema auditivo não tem capacidade para distinguir um tom de outro

Nº Banda Crítica	Frequência Corte Inferior (Hz)	Frequência Central (Hz)	Frequência Corte Superior (Hz)	Banda Crítica (Hz)
1	---	50	100	----
2	100	150	200	100
3	200	250	300	100
4	300	350	400	100
5	400	450	510	110
6	510	570	630	120
7	630	700	770	140
8	770	840	920	150
9	920	1000	1080	160
10	1080	1170	1270	190
11	1270	1370	1480	210
12	1480	1600	1720	240
13	1720	1850	2000	280
14	2000	2150	2320	320
15	2320	2500	2700	380
16	2700	2900	3150	450
17	3150	3400	3700	550
18	3700	4000	4400	700
19	4400	4800	5300	900
20	5300	5800	6400	1100
21	6400	7000	7700	1300
22	7700	8500	9500	1800
23	9500	10500	12000	2500
24	12000	13500	15500	3500
25	15500	18775	22050	6550

## Efeito de máscara temporal

- Um tom máscara outro que ocorra dentro de uma janela temporal:
  - Quanto maior o nível de som, maior a duração da máscara
  - Depende da duração do tom aplicado



- Máscara no tempo e na frequência

